Bivariate Multi-Fractal Model: Estimation of parameters and Applications to Risk Management

R. Liu and T. Lux

Department of Economics
University of Kiel, Germany

Abstract

Long memory (long-term dependence) seems to be as widespread in financial time series as in nature. Inspired by the long memory property, multi-fractal processes have recently been introduced as a new tool for modeling the stylized facts in financial time series. In this paper, we attempt to construct bivariate multi-fractal model, and implement its estimation via both GMM and likelihood approaches. For its empirical assessment, we apply the model on portfolio investment concerning VaR using time series of stock exchange indices, foreign exchange rates and bond maturity rates.

Keyword: Long memory, Bivariate Multifractal, GMM estimation, VaR.

JEL Classification: C20, G15

Correspondence: Department of Economics, University of Kiel, Olshausenstr. 40; 24118 Kiel, Germany.

1 Introduction

Following statistical analyses like Hurst's R/S test and the modified R/S by Lo (1991); as well as some econometric models such as ARFIMA (Fractional Integrated Autoregressive Moving Average), and FIGARCH (Fractional Integrated General Autoregressive Conditional Heteroscedasticity), the Multi-Fractal Model (MF) has been recently introduced as an alterative formalisation, which conceives volatility as a hierarchical, multiplicative process with heterogeneous components. The essential new feature of MF models is their ability of generating different degrees of long-term dependence in various powers of returns - a feature pervasively found in empirical financial data (Ding et al 1993). Research on Multi-Fractal models originated from statistical physics (Mandelbrot, 1974). Unfortunately, the models used in physics are of a combinatorial nature and suffer from non-stationarity due to the limitation to a bounded interval and the non-convergence of moments in the continuous-time limit. This major weakness was overcome by introducing a iterative version of the multi-fractal model.

So far, available multi-fractal models are mostly univariate ones [1]. However, for many important questions in empirical research, multi-variate settings are preferable. In particulary it is now well accepted that financial volatilities move together over time across assets and markets. This is particularly important when considering asset allocation, value-at-risk and portfolio hedging strategies. Secondly, since the information on the source of long memory in the volatility process is quite limited, the Multivariate (bivariate) model may provide additional insight into the factors responsible for long memory.

The rest of this paper is organized as follows: Section 2 provides a review of the Multi-fractal model of financial returns. Section 3 introduces the bivariate multi-fractal model and different approaches to its estimation. For GMM, we give the details of the analytical moments in the Appendix. Empirical works on the base of this new bi-variate process is presented in Section 4.

2 Review of Multifractal Model

Financial markets display some similarities to fluid turbulence. For example, both turbulent fluctuations and financial fluctuations display intermittency at all scales. A cascade of energy flux is known to occur from the large scale of injection to the small scale of dissipation. This cascade is typically modeled by a multiplicative cascade, which then leads to multi-fractal field.

Mandelbrot et al. (1997) first introduced the Multi-Fractal Model, translating the approach of Mandelbrot (1974) from the statistical physics area into finance. Fisher and Calvet (2004a) report advantages of Multi-Fractal models compared to GARCH and FIGARCH in various financial time series. Lux (2004b) provides related evidence on forecasting of future volatility generated from the Multi-Fractal model, the results demonstrating its potential advantage.

The first type of the MF proposed by Mandelbrot et al. (1997), named the Multi-Fractal Model of Assets Returns (MMAR), assumes that returns $x(t)$ follow a compound process:

$$x(t) = B_H[\theta(t)] \tag{1}$$

in which an incremental fractional Brownian motion with index H, $B_H[\cdot]$, is subordinate to the cumulative distribution function $\theta(t)$ of a multi-fractal measure, which was already employed by Mandelbrot (1974), when modeling the distribution of energy in turbulent dissipation.

[1] An exception is Calvet, et al (2004b) whose approach, however, differs from ours in various aspects.

The simplest way to create a multi-fractal measure is the "binomial multi-fractal", constructed on a unit interval [0,1] with uniform density. one proceeds as follows: Divide the interval into two subintervals of equal length. Let m_0 and m_1 be two positive numbers adding up to 1. In the first step, this interval is split into two equal subintervals, and the measure uniformly spreads mass equal to m_0 on the subinterval [0, 0.5] and mass equal to m_1 on [0.5, 1], in step 2, the set [0, 0.5] is split into two subintervals, [0, 0.25] and [0.25, 0.5]; which respectively receive a fraction measure m_0 and m_1 of the total mass [0, 0.5]; We apply the same procedure to the dyadic set [0.5, 1], and the above procedure is then repeated ad infinitum, and iteration of this procedure generates an infinite sequence of measures.

As a minor extension of the original binomial measure one could simply dispense with the rule of always assigning m_0 to the left, and m_1 to the right, randomizing the assignment instead; or, one may uniformly split the interval into an arbitrary number b larger than 2 at each stage of the cascade, and receive the fractions $m_0, m_1 \dots m_{b-1}$, which leads to a so-called multinomial measure. Furthermore, we can also randomize the allocations between the subintervals, taking $m_0, m_1 \dots m_{b-1}$ with certain probabilities, or using random numbers for m_0 instead of the same constant value, such as draws from a Lognormal distribution in Mandelbrot (1974, 1997).

The above mechanism may be called a combinatorial MF model. It is immediately obvious that one important limitation of this approach is the limited domain of resulting measure. With an underlying cascade extending over k steps, we have exactly 2^k (Binary cascade) different subintervals at our disposal and, therefore, could generate only "time series" which are no longer than 2^k. Later, this difficulty was overcome by the introduction of an iterative Markov-switching MF model in Calvet and Fisher (2001). In their approach, returns are modeled as:

$$x_t = \sigma \left(\prod_{i=1}^{k} M_t^{(i)} \right)^{1/2} \cdot u_t \tag{2}$$

with u_t drawn from a standard Normal distribution $N(0,1)$ and instantaneous volatility being determined by the product of k volatility components or multipliers $M_t^{(1)}$, $M_t^{(2)}$ ..., $M_t^{(k)}$, [2] and a constant scale parameter σ.

Each volatility component is renewed at time t with probability γ_i depending on its rank within the hierarchy of multipliers or remains unchanged with probability $1 - \gamma_i$. The transition probabilities are specified as:

$$\gamma_i = 1 - (1 - \gamma_1)^{(b^{k-1})} \tag{3}$$

with parameters $\gamma_1 \in [0, 1]$ and $b \in (1, \infty)$. Estimation of this model, then,

[2] Additionally, $E[M_t^{(i)}]$ or $E[\sum M_t^{(i)}]$ equal to some arbitrary value is usually imposed for the sake of normalizing the time-varying components of volatility. Both Calvet and Fisher (2002) and Lux (2004) assume a Binomial distribution with parameters m_0 and $2 - m_0$ (thus, guaranteeing an expectation of unity for all $M_t^{(i)}$), and for a Lognormal distribution $E[M] = 1$.

involves the parameters γ_1 and b as well as those characterizing the distribution of the components $M_{i,t}$.

The main attraction of Multi-Fractal model is that it shares certain properties of asset returns: fat tails and asymptotic power-law behavior of the autocovariance function (long memory)[3] . Furthermore, multifractality implies that different powers of the measure have different decay rates of their autocovariances. Calvet and Fisher (2002) show that this feature carries over to absolute moments of returns in the MMAR (eq.1). In this sense, other alternatives like FIGARCH or ARFIMA models belong to the catalogue of uni-fractal model, i.e. they have the same decay rate for all moments. Although the Multi-Fractal model is a rather new model in financial economics, there are already various attempts at estimating the parameters of the multi-fractal model. Available options include the traditional Scaling estimator; GMM estimation introduced by Lux (2003, 2004b) and Maximum Likelihood Estimation derived by Calvet and Fisher (2004a), which will both be used in the next section.

3 The Bivariate Multi-Fractal Model and its Estimation

We introduced a parsimonious Bivariate Multi-fractal model (BMF) under the hypothesis of two time series having certain amounts of joint cascade levels in common in both multi-fractal processes.

$$r_{q,t} = \sigma_q \cdot \left[\left(\prod_{i=1}^{k} M_t^{(i)} \right) \cdot \left(\prod_{l=k+1}^{n} M_t^{(l)} \right) \right]^{1/2} \cdot u_{q,t} \tag{4}$$

$q = 1, 2$ refers to two time series, both having an overall number of n levels of their volatility cascades, and they share k numbers of joint cascade levels which govern the strength of their volatility correlation. σ_q are the unconditional standard deviation of the return series. Obviously, the larger k, the more correlation between them. After k joint multiplications, each series has separate additional multifractal components. The increments $u_{q,t}$ follow a bivariate standard Normal distribution with correlation parameter ρ.[4]

Furthermore, we restrict the specification of the transition probabilities to:

$$\gamma_j = 2^{-(k'-j)} \tag{5}$$

Each component is renewed at time t with probability γ_i depending on its rank within the hierarchy of multipliers and remains unchanged with probability $1 - \gamma_i$.

[3] I.e.$Cov(|x_t|^q, |x_{t+\tau}|^q) \propto \tau^{2d(q)-1}$. However, one should note that the Markov-switching multi-fractal of eq. (2) only has "long memory over a limited range", cf. Calvet and Fisher, 2001, for details.

[4] The independence of increments was assumed in former version of this paper.

We specify the multipliers to be random draws from either a Binomial or Lognormal distribution, In the binomial case in which we assume two draws $m_0 \in (0, 2)$ and alternative $m_1 = 2 - m_0$, for the latter, we assume $-log_2 M \sim N(\lambda, \sigma^2)$, and assign constraint $E[M_t^{(i)}] = 0.5$ which leads to $\sigma_m^2 = 2(\lambda - 1)/ln2$.

Figure 1 and Figure 2 show simulations of the bivariate multi-fractal model (k = 4, n = 20) with Binomial distribution of its multipliers together with its ACFs. The simulation apparently shares some of the stylized facts of financial time series, namely volatility clustering and hyperbolical decay of the autocorrelation function. One also easily recognizes the correlation in the volatility of both time series.

3.1 Generalized Method of Moments Estimation

Historically, the first attempt at estimating the multi-fractal models is the scaling estimator. Since multifractal measures are characterized by a non-linear scaling function of moments (scaling law), through a Legendre transformation, parameter estimation is achieved by matching the empirical and hypothetical spectrum of Hölder exponents. In our proceeding bivariate MF model, we will, however, exclude the scaling estimator due to its bias and lack of asymptotic distribution theory, cf. Lux (2003, 2004a).

Instead, we adopt the GMM (Generalized Method of Moments) approach by Hansen (1982) with analytical solutions of a set of appropriate moment conditions. In the GMM approach, the vector of parameter estimates of the model, say β, can be obtained as:

$$\widehat{\beta} = \arg\min_{\beta \in \Theta} \bar{M}(\beta)' W \bar{M}(\beta) \tag{6}$$

with β the parameter vector, $\bar{M}(\beta)$ the vector of differences between sample moments and analytical moments, and W a positive definite weighting matrix, which controls the over-identification when applying GMM. Implementing (6), one typically starts with the identity matrix, then the inverse of the covariance matrix obtained from the first round estimation is used as the weighting matrix in the next step, and the procedure will continue until the estimates and weighting matrices converge. Under suitable conditions, $\hat{\beta}$ is consistent and asymptotically converges to $T^{1/2}(\hat{\beta} - \beta_0) \sim N(0, \Xi)$ with covariance matrix Ξ.

The applicability of GMM for multi-fractal models has been discussed by Lux (2003). The approach recommended in this paper is using log differences of absolute returns together with the pertinent analytical moment conditions, i.e.

to transform the observed data r_t into Tth differences of the log observations:

$$\begin{aligned}
X_{t,T} &= \ln|r_{1,t}| - \ln|r_{1,t-T}| \\
&= \left(0.5\sum_{i=1}^{k}\varepsilon_t^{(i)} + 0.5\sum_{h=k+1}^{n}\varepsilon_t^{(h)} + ln|u_t|\right) - \left(0.5\sum_{i=1}^{k}\varepsilon_{t-T}^{(i)} + 0.5\sum_{h=k+1}^{n}\varepsilon_{t-T}^{(h)} + ln|u_{t-T}|\right) \\
&= 0.5\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-T}^{(i)}) + 0.5\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-T}^{(h)}) + (ln|u_t| - ln|u_{t-T}|)
\end{aligned} \tag{7}$$

with $\varepsilon_t^{(i)} = \ln\left(M_t^{(i)}\right)$, and in the same way to define the second time series, say $Y_{t.T}$.

In order to exploit as much as possible information, the moment conditions that we consider include two categories: the first set of conditions is obtained by considering some order of log-squared observations, and the second set of moment conditions is derived from the absolute observations. In particular, we select moment conditions for the powers of $X_{t,T}$ and $Y_{t,T}$, i.e. moments of the raw observations and square observations:

$$Cov[X_{t,1}^q, Y_{t,1}^q];\ Cov[X_{t+1,1}^q, Y_{t,1}^q];\ Cov[X_{t+T,T}^q, X_{t,T}^q];\ Cov[Y_{t+T,T}^q, Y_{t,T}^q]$$

for $q = 1, 2$ and $T = 1, 5, 10, 20$. It is straightforward to get the moments for the raw observations, but the moment calculations for the squared data seem a bit tedious. The detailed analytical moments are given in the Appendix.

We proceed by conducting several Monte Carlo experiments to explore the performance of the GMM estimation. Moment conditions for the Binomial and Lognormal distribution can be found in Appendices A and B. We start with the Binomial Model (n = 20) with number of joint multipliers $k = 4$ and $k = 8$, we fixed correlation parameter $\rho = 0.5$, and choose multipliers from $m_0 = 1.2$ to 1.5 by 0.1 increment with sample sizes $N_1 = 2000$, $N_2 = 5000$, and $N_3 = 10000$. Table 1 shows the statistical result of the GMM estimator: for the Binomial distribution parameter $\hat{m_0}$, not only the Bias but also the finite sample standard deviation and root mean squared error show quite encouraging behavior, even in the small sample size $N = 2000$ and $N = 5000$, the average bias of the Monte Carlo estimates is moderate throughout and practically zero for the larger sample sizes $N = 10000$.

It is also interesting to note that our estimates are in harmony with $T^{\frac{1}{2}}$ consistency, and the Hansen's J test reveals that there is not disappointing concerning the over-identification restrictions (see Figure 3). All these results can be viewed as a positive signal of the log transformation in practice. Furthermore, we also notice that there is almost no significant difference between $k = 4$ and $k = 8$ in Table 1, the very slight sensitivity of the estimates of m_0 with respect to the number of joint cascades might be viewed as a very welcome phenomenon as it implies that estimation of m_0 is hardly affected by the potential mis-specification of the number of joint cascade k.

Then, we turn to the Bivariate MF with a continuous distribution ($-log_2 M \sim N(\lambda, \sigma^2)$). In our Monte Carlo simulations reported in Table 2, we cover parameter values $\lambda = 1.10$, 1.20 1.30 and 1.40, and use the same numbers of joint multiplier cascade levels and the sample sizes as in the Binomial case above. As can be seen, results are not too different from those obtained with the Binomial model: Biases are moderate again, and results for λ are almost insensitive with respect to k. Somewhat in contrast to the Binomial case, we notice a very slight deterioration of efficiency with increasing λ, which might be due to increasing λ leading to increasing σ_m^2 by their dependence (recalling that $\sigma_m^2 = 2(\lambda - 1)/ln2$). All in all, the results from both the Binomial and Lognormal Monte Carlo simulation and estimation show that GMM seems to work quite well for multi-fractal processes both in the discrete and in the continuous state space.

3.2 Maximum Likelihood Estimation

The MF dynamics can be interpreted as a special case of a Markov-switching process with a large state space. This makes Maximum Likelihood Estimation feasible. In our parsimonious bivariate MF model, the state spaces is finite when the multipliers follow a discrete distribution (i.e. Binomial distribution). The likelihood function can be derived by determining the exact form of each possible component in the transition matrix, and is similar to the likelihood function developed for the uni-variate process by Calvet et al (2004b), but differs in so far as the transition matrix of each multifractal component contains two starting cascade level. r_t is the set of joint return observations $\{r_{q,t}\}$ for $q =$ 1,2. We have the likelihood function below:

$$\begin{aligned} f(r_1, \cdots, r_T; \Theta) &= \prod_{t=1}^{T} f(r_t | r_1, \cdots, r_{t-1}) \\ &= \prod_{t=1}^{T} \left[\sum_{i=1}^{4^n} P(M_t = m^i | r_1, \cdots, r_{t-1}) \cdot f(r_t | M_t = m^i) \right] \\ &= \prod_{t=1}^{T} (\Omega_{t-1} A) \cdot f(r_t | M_t = m^i). \end{aligned} \tag{8}$$

With transition matrix A which has components A_{ij} equal to

$$\begin{aligned} & P(M_{t+1} = m^j | M_t = m^i) \\ = & \prod_{k=1}^{n} \left[(1 - \gamma_k) \cdot 1_{\{m_k^i = m_k^j\}} + \gamma_k P(M_t = m_k^j) \right] \end{aligned} \tag{9}$$

Both M_t and $m^{(\cdot)}$ are vectors, $M_t = (M_t^1, \cdots, M_t^k, M_t^{k+1}, \cdots, M_t^n)$, m_k^i denotes the kth component of vector m^i.

The density of the innovation r_t conditional on M_t is:

$$f(r_t|M_t = m^i) = \frac{F_N\left\{r_t/\left[\sigma\cdot\left(\prod_{i=1}^{k} M_t^{(i)} \prod_{j=k+1}^{n} M_t^{(j)}\right)^{0.5}\right]\right\}}{\sigma\cdot\left(\prod_{i=1}^{k} M_t^{(i)} \prod_{j=k+1}^{n} M_t^{(j)}\right)^{0.5}} \tag{10}$$

$F_N\{\cdot\}$ denotes the bivariate Normal density function.

The last unknown component in the likelihood function above is Ω, which is the conditional probability defined by $\Omega_t^i = P(M_t = m^i|r_1, \cdots, r_t)$, and due to $\sum_{i=1}^{4^n} \Omega_t^i = 1$, by Bayesian updating, we get[5]

$$\Omega_{t+1} = \frac{f(r_t|M_t = m^i) \bigotimes (\Omega_t A)}{\sum f(r_t|M_t = m^i) \bigotimes (\Omega_t A)} \tag{11}$$

It would be not too surprising that the ML estimators are more efficient compared with the two previous tables, as ML extracts all the information in the data. However, applicability of the ML approach is constrained by its computational demands: First, it is not applicable for models with an infinite state space, i.e. continuous distributions of the multipliers such as Lognormal distribution we use here. Secondly, even for the discrete distributions, say the Binomial case, current computational limitations make choices of cascades with a number of steps n beyond 5 unfeasible because of the implied evaluation of a $4^n \times 4^n$ transition matrix in each iteration. Table 3 presents the comparison of ML and GMM estimator (Monte Carlo design as previous tables) in the case of $n = 5$, $k = 2$. Table 4 reports the performance of ML estimator including the scaling parameters.

3.3 Simulated based Maximum Likelihood Estimation

Particle filter.

4 Value at Risk

One widely used tool to gear and control market risk is Value-at-Risk (VaR), which measures the worst loss over a specified target horizon with a given statistical confidence level. In other words, it represents a quantile of an estimated profit-loss distribution. Various organizations and interest groups have recommended VaR as a portfolio risk-measurement tool.

[5] $\bigotimes$ represents element by element product.

In this section, We will report empirical work on the application of the Bivariate MF model for value at risk assessment. $\tilde{r}_{t,t+h}$ is defined as the forward-looking h-period return at time t: $\tilde{r}_{t,t+h} = \sum_{i=1}^{h} r_{t+i}$. VaR at the h-period horizon is defined as the α quantile of the conditional probability distribution of $\tilde{r}_{t:t+h}$:

$$Pr\left(\tilde{r}_{t:t+h} \leq VaR^{\alpha}_{t:t+h}|I_t\right) = \alpha. \tag{12}$$

The performance of our model is assessed by computing the failure rate for the returns and the portfolio. By definition, the failure rate is the number of times returns exceed (here in absolute value) the forecasted VaR. If the model is well specified, the failure rate is expected to be as close as possible to the prespecified VaR level.

In the empirical application we consider daily data for a collection of Stock Exchange Index: Dow Jones Composite 65 Average Index and NIKKEI 225 Average Index (DOW/NIK, Jan. 1969 - Aug. 1999); two Foreign Exchange rates: British Pound to US Dollar , Australian Dollar to US Dollar (BP/AUD, March 1973 - Nov. 2004); and U.S. 1 Year and 2 Year Treasury Constant Maturity Bond Rate (T1/T2, June 1976 - Oct.2004), where the first symbol inside the parentheses designates the short notation for the time series and the numbers in parentheses are the start and end period for the sample at hand. For all time series daily observations are denoted p_t, returns are defined as $r_t = ln(p_t) - ln(p_{t-1})$.

Empirical results for the time series are given in tables 5 and 6. We estimate the bivariate MF model by GMM and Maximum Likelihood. In GMM estimation, we employ both the Binomial model (for Stocks and Bonds) and with the Lognormal model (for Foreign Exchange rate). We then simulate the bivariate time series based on the estimators, and calculate forward-looking h-period returns for single time series and portfolios. $VaR_{t,h}(p)$ is obtained as the $(1-p)^{th}$ empirical quantile, we calculate the failure rate to assess the performance of our bivariate MF model, if the model is well specified, the failure rate should be close to the pre-assume confident level. We describe the results with respect each pair time series as following:

(1) For Stock Exchange Index DOW and NIK, the VaR forecasts based on GMM and ML are quite satisfactory, except with only one too risky case for individual Dow Jones index at 5% confident level .

(2) For Foreign Exchange rates BP and AUD, a remarkable result we have is that VaR by GMM estimation (table 5) is very successful throughout all cases. On the contrary, table 6 (VaR based ML estimator) reports some conservative VaRs for Equal Weighting portfolio: one in 1-day horizon, and two cases in 2-day and 5-day horizon. We recognize that excessive conservativeness does not imply superior risk management in financial investment.

(3) For US Bond maturity rates, we first find the success at 10% level from both table, furthermore for T1, VaR forecasts from both GMM and ML are also well specified in all horizons. In contrast, for T2 and Equal Weighting portfolio (EW), it leaves several too risky VaRs in both table 5 and table 6 at confidence levels of 1% and 5%, which is of course against the investment principle.

5 Conclusion

In this paper we have developed a bivariate Multi-Fractal model extending the univariate Markov-switching Multi-Fractal model, and we implemented both GMM and Maximum Likelihood estimation. For GMM, eight moments conditions have been employed through the log transformation of observations. Our Monte Carlo experiments indicate the positive performance of both GMM and ML estimators. Although GMM is not as efficient as ML, it has the important advantage that it does not pose computational restrictions on the choice of the number of cascade levels with GMM compared to a maximum of about 5 cascade levels in ML estimation. Furthermore, empirically speaking, GMM is much faster compared to the very time-consuming ML process.

In the last part of this paper, we applied the model to Value-at-Risk assessment with empirical financial time series of Stock Exchange index, Foreign Exchange rates and Bond maturity rates. We demonstrate the applicability of the bivariate Multi-Fractal model, and the results also present that GMM estimator is competitive and could generate better VaR forecasts in some cases (here exchange rate case).

References

Baillie et al (1996) Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, **74**, 3-30.

Beran, J. (1994) *Statistics for Long-Memory Processes*, New York: Chapman & Hall.

Barkoulas. J.T and Baum, C.F (2000) Long Memory in the Greek Stock Market, *Applied Financial Economics*, **10**, 177-184.

Calvet, L. and A. Fisher (2001) Forecasting Multifractal Volatility, *Journal of Econometrics*, **105**, 27-58.

Calvet, L. and A. Fisher (2002) Multi-Fractality in Asset Returns: Theory and Evidence, *Review of Economics & Statistics*, **84**, 381-406.

Calvet, L. and A. Fisher (2004a) Regime Switching and the Estimation of Multifractal Processes, *Journal of Financial Econometrics*, **2**,49-83.

Calvet, et al (2004b) *Volatility Comovement: A Multifrequency Approach*, NBER Technical Working Paper 0300.

Cavalcante, Jorge and Assaf, A. (2002) *Long Range Dependence in the Returns and Volatility of the Brazilian Stock Market* Manuscript, Rio de Janeiro.

Chung, C.-F. (2002) *Estimating the Fractionally Integrated GARCH Model*, Mimeo: Academia Sinica.

Diebold, F. and R. Mariano (1995) Comparing Predictive Accurancy, *Journal of Business and Economics Statistics*, **13**,253-263.

Ding et al (1993) A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance*, **1**, 83-106.

Evertz, C.J.G. and B. Mandelbrot (1992) Multifractal Measures, in: Peitgen, H.-O., H. Jrgens and D. Saupe, eds., *Chaos and Fractals: New Frontiers of Science*, Berlin: Springer.

Golaka. C Nath (2002) *Long Memory and Indian Stock Market - an Empirical Evidence* Manuscript, WDM Dept, Mumbai.

Hansen, L.P (1982) Large Sample Properties of Generalized Method of Moments

Estimators, *Econometrica*, **50**, 1029-1054.

Harte, D. (2001) Multifractals: Theory and Applications, London: Chapman and Hall.

Lo. A.W (1991): Long-Term Memory in Stock Market Prices, *Econometrica*,**59**, 1279-1313.

Lobato. I. N. and N.E. Savin (1998) Real and Spurious Long-Memory Properties of Stock Market Data, *Journal of Business* & *Economics Statistics*, **16**, 261-283.

Lux, T. (1996) Long-term Stochastic Dependence in Financial Prices: Evidence from the German Stock Market, *Applied Economics Letters*, **3**, 701-706.

Lux, T. (2003) *The Multi-Fractal Model of Asset Returns, Its Estimation via GMM and Its Use for Volatility Forecasting* University of Kiel.

Lux, T. (2004a) Detecting Multi-Fractal properties in Asset Returns: The Failure of the "Scaling Estimator" *International Journal of Modern Physics*, **15**, 481-491.

Lux, T. (2004b) *The Markov-Switching Multi-Fractal Model of Asset Returns: GMM Estimation and Linear Forecasting of Volatility*, University of Kiel.

Madhusoodanan, T. P (1998): Long-Term Dependence in the Indian Stock Market. *Journal of Financial Studies*, 5, 33-53.

Mandelbrot, B., (1974) Intermittent Turbulence in Self Similar Cascades: Divergence of High Moments and Dimension of Carrier, *Journal of Fluid Mechanics*, **62**, 331-358.

Mandelbrot et al (1997) *A Multifractal Model of Asset Returns*, Mimeo: Cowles Foundation for Research in Economics.

Tolvi. J. (2003): Long-Term Memory in a Small Stock Market, *Economics Bulletin*, **7**, 1-13.

Zumbach. G (2002) *Volatility Processes and Volatility Forecast with Long Memory*, Manuscript, Zürich.

Appendix: Moment Conditions

Recall the model from Section 3. Let $\varepsilon_t^{(\cdot)} = ln(|M_t^{(\cdot)}|)$, and we compute the first log difference:

$$
\begin{aligned}
X_{t,1} &= ln(|r_{1,t}|) - ln(|r_{1,t-1}|) \\
&= \left(\frac{1}{2}\sum_{i=1}^{k}\varepsilon_t^{(i)} + \frac{1}{2}\sum_{l=k+1}^{n}\varepsilon_t^{(l)} + ln|u_{1,t}|\right) - \left(\frac{1}{2}\sum_{i=1}^{k}\varepsilon_{t-1}^{(i)} + \frac{1}{2}\sum_{l=k+1}^{n}\varepsilon_{t-1}^{(l)} + ln|u_{1,t-1}|\right) \\
&= \frac{1}{2}\sum_{i=1}^{k}\left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right) + \frac{1}{2}\sum_{l=k+1}^{n}\left(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}\right) + (ln|u_{2,t}| - ln|u_{1,t-1}|)
\end{aligned}
$$

$$
\begin{aligned}
Y_{t,1} &= ln(|r_{2,t}|) - ln(|r_{2,t-1}|) \\
&= \left(\frac{1}{2}\sum_{i=1}^{k}\varepsilon_t^{(i)} + \frac{1}{2}\sum_{h=k+1}^{n}\varepsilon_t^{(h)} + ln|u_{2,t}|\right) - \left(\frac{1}{2}\sum_{i=1}^{k}\varepsilon_{t-1}^{(i)} + \frac{1}{2}\sum_{h=k+1}^{n}\varepsilon_{t-1}^{(h)} + ln|u_{2,t-1}|\right) \\
&= \frac{1}{2}\sum_{i=1}^{k}\left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right) + \frac{1}{2}\sum_{h=k+1}^{n}\left(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}\right) + (ln|u_{2,t}| - ln|u_{2,t-1}|)
\end{aligned}
$$

A Binomial case

$$
\begin{aligned}
& cov[X_{t,1}, Y_{t,1}] \\
&= E\left[(X_{t,1} - E[X_{t,1}]) \cdot (Y_{t,1} - E[Y_{t,1}])\right] = E[X_{t,1} \cdot Y_{t,1}] \\
&= E\left\{\left[\frac{1}{2}\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \frac{1}{2}\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}) + (ln|u_{1,t}| - ln|u_{1,t-1}|)\right] \cdot \right. \\
&\qquad \left.\left[\frac{1}{2}\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \frac{1}{2}\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}) + (ln|u_{2,t}| - ln|u_{2,t-1}|)\right]\right\} \\
&= \frac{1}{4}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right] - 2E[u_t]^2 + 2E[ln|u_{1,t}| \cdot ln|u_{2,t}|].
\end{aligned}
\tag{A1}
$$

We firstly consider $E[(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})^2]$, the only one non-zero contribution is $[ln(m_0) - ln(2 - m_0)]^2$, and it occurs when new draws take place in cascade level i between t and $t-1$, whose probability by definition is $\frac{1}{2}\frac{1}{2^{k-i}}$. Summing up we get:

$cov[X_{t,1}, Y_{t,1}] = 0.25\cdot[ln(m_0)-ln(2-m_0)]^2\cdot\sum_{i=1}^{k}\frac{1}{2}\frac{1}{2^{k-i}}-2E[u_t]^2+2E[ln|u_{1,t}|\cdot ln|u_{2,t}|]$.

$$
\begin{aligned}
&cov[X_{t+1,1}, Y_{t,1}] \\
&= E\left\{\left[\frac{1}{2}\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})+\frac{1}{2}\sum_{l=k+1}^{n}(\varepsilon_{t+1}^{(l)}-\varepsilon_{t}^{(l)})+(ln|u_{1,t+1}|-ln|u_{1,t}|)\right]\cdot\right. \\
&\quad\left.\left[\frac{1}{2}\sum_{i=1}^{k}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})+\frac{1}{2}\sum_{h=k+1}^{n}(\varepsilon_{t}^{(h)}-\varepsilon_{t-1}^{(h)})+(ln|u_{2,t}|-ln|u_{2,t-1}|)\right]\right\} \\
&= \frac{1}{4}E\left[\sum_{i=1}^{k}\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)\cdot\sum_{i=1}^{k}\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)\right]+E[u_t]^2-E[ln|u_{1,t}|\cdot ln|u_{2,t}|].
\end{aligned}
\tag{A2}
$$

For $(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})$, the non-zero value only occurs in case of two changes of the multiplier from time $t+1$ to time $t-1$, the probability of this occurrence is $(\frac{1}{2}\frac{1}{2^{k-i}})^2$. So, we have the result:

$$
\begin{aligned}
&cov[X_{t+1,1}, Y_{t,1}] \\
&= 0.25\cdot[2ln(m_0)\cdot ln(2-m_0)-(ln(m_0))^2-(ln(2-m_0))^2]\cdot\sum_{i=1}^{k}(\frac{1}{2}\frac{1}{2^{k-i}})^2 \\
&+E[u_t]^2-E[ln|u_{1,t}|\cdot ln|u_{2,t}|].
\end{aligned}
$$

Then, we look at the moment condition for one single time series:

$$
\begin{aligned}
&cov[X_{t+1,1}, X_{t,1}] \\
&= E\left\{\left[\frac{1}{2}\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})+\frac{1}{2}\sum_{l=k+1}^{n}(\varepsilon_{t+1}^{(l)}-\varepsilon_{t}^{(l)})+(ln|u_{1,t+1}|-ln|u_{1,t}|)\right]\cdot\right. \\
&\quad\left.\left[\frac{1}{2}\sum_{i=1}^{k}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})+\frac{1}{2}\sum_{l=k+1}^{n}(\varepsilon_{t}^{(l)}-\varepsilon_{t-1}^{(l)})+(ln|u_{1,t}|-ln|u_{1,t-1}|)\right]\right\} \\
&= \frac{1}{4}E\left[\sum_{i=1}^{k}\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)\cdot\sum_{i=1}^{k}\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)\right]+\frac{1}{4}E\left[\sum_{l=k+1}^{n}\left(\varepsilon_{t+1}^{(l)}-\varepsilon_{t}^{(l)}\right)\cdot\sum_{l=k+1}^{n}\left(\varepsilon_{t}^{(l)}-\varepsilon_{t-1}^{(l)}\right)\right] \\
&+E\left[ln|u_t|\right]^2-E\left[ln|u_t|^2\right].
\end{aligned}
\tag{A3}
$$

The first component is identical to the one of the case of $cov[X_{t+1,1}, Y_{t,1}]$, and the second component can be derived in the same way. Adding together we arrive at:

$$
\begin{aligned}
&cov[X_{t+1,1}, X_{t,1}] \\
&= 0.25 \cdot \left[2ln(m_0) \cdot ln(2-m_0) - (ln(m_0))^2 - (ln(2-m_0))^2\right] \cdot \sum_{i=1}^{k} (\tfrac{1}{2} \tfrac{1}{2^{k-i}})^2 \\
&+0.25 \cdot \left[2ln(m_0) \cdot ln(2-m_0) - (ln(m_0))^2 - (ln(2-m_0))^2\right] \cdot \sum_{i=k+1}^{n} (\tfrac{1}{2} \tfrac{1}{2^{n-i}})^2 \\
&+E[ln|u_t|]^2 - E[ln|u_t|^2].
\end{aligned}
\tag{A4}
$$

By our assumption of both time series having the same number of cascade levels, the moments for the two individual time series are identical for the same length of time lags.

Then, let's turn to the squared observations:

$$
\begin{aligned}
&E[X_{t,1}^2 \cdot Y_{t,1}^2] \\
&= E\left\{\left[\tfrac{1}{2}\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \tfrac{1}{2}\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}) + (ln|u_{1,t}| - ln|u_{1,t-1}|)\right]^2 \cdot \right. \\
&\qquad \left. \left[\tfrac{1}{2}\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \tfrac{1}{2}\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}) + (ln|u_{2,t}| - ln|u_{2,t-1}|)\right]^2 \right\} \\
&= \tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^4\right] + \tfrac{1}{16}E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2 \left(\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)})\right)^2\right] \\
&+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2 \left(\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)})\right)^2\right] \\
&+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2 \left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\right] \\
&+\tfrac{1}{4}\left\{2E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right] + 2E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\right]\right\} \cdot \left(2E[ln|u_t|^2] - 2E[ln|u_t|]^2\right) \\
&+2E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 8E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|] \\
&+4E[ln|u_{1,t}| \cdot (ln|u_{2,t}|]^2 + 2E[(ln|u_t|)^2]^2.
\end{aligned}
$$

By examining each component in the expression above combining with the calculations of the previous moments, it is not difficult to find the solution:

$E[X_{t,1}^2 \cdot Y_{t,1}^2]$

$= [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \frac{1}{16} \sum_{i=1}^{k} \frac{1}{2} \frac{1}{2^{k-i}} + [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \frac{1}{16} \sum_{i=k+1}^{n} \frac{1}{2} \frac{1}{2^{n-i}} \sum_{i=k+1}^{n} \frac{1}{2} \frac{1}{2^{n-i}}$

$+2\,[ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \frac{1}{16} \sum_{i=1}^{k} \frac{1}{2} \frac{1}{2^{k-i}} \sum_{i=k+1}^{n} \frac{1}{2} \frac{1}{2^{n-i}}$

$+ \left(E[ln|u_t|^2] - E[ln|u_t|]^2\right) \cdot [ln\,(m_0) - ln\,(2-m_0)]^2 \cdot \left(\sum_{i=1}^{k} \frac{1}{2} \frac{1}{2^{k-i}} + \sum_{i=k+1}^{n} \frac{1}{2} \frac{1}{2^{n-i}} \right)$

$+2E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 8E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|]$

$+4E[ln|u_{1,t}| \cdot ln|u_{2,t}|]^2 + 2E[(ln|u_t|)^2]^2.$

(A5)

$E[X_{t+1,1}^2 \cdot Y_{t,1}^2]$

$= E\left\{ \left[\frac{1}{2} \sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) + \frac{1}{2} \sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) + (ln|u_{1,t+1}| - ln|u_{1,t}|) \right]^2 \cdot \right.$

$\left. \left[\frac{1}{2} \sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \frac{1}{2} \sum_{h=k+1}^{n} (\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}) + (ln|u_{2,t}| - ln|u_{2,t-1}|) \right]^2 \right\}$

$= \frac{1}{16} E\left[\left(\sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \left(\sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \right]$

$+ \frac{1}{16} E\left[\left(\sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) \right)^2 \left(\sum_{h=k+1}^{n} (\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}) \right)^2 \right]$

$+ \frac{1}{16} E\left[\left(\sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \left(\sum_{h=k+1}^{n} (\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}) \right)^2 \right]$

$+ \frac{1}{16} E\left[\left(\sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \left(\sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) \right)^2 \right]$

$+ \frac{1}{4} \left\{ 2E\left[\left(\sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \right] + 2E\left[\left(\sum_{l=k+1}^{n} (\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}) \right)^2 \right] \right\} \cdot \left(2E[ln|u_t|^2] - 2E[ln|u_t|]^2\right)$

$+E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 4E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|] + 4E[ln|u_{1,t}| \cdot ln|u_{2,t}|]E[ln|u_t|]^2$

$+3E[ln|u_t|^2]^2 - 4E[ln|u_t|^2]E[ln|u_t|]^2.$

Until now, the only unfamiliar component is the first term:

$E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})\right)^{2}\cdot\left(\sum_{i=1}^{k}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})\right)^{2}\right]$, there are three different forms to be considered:

(1) $\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)^{2}\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)^{2}$, which have non-zero value only if $\varepsilon_{t+1}^{(i)} \neq \varepsilon_{t}^{(i)} \neq \varepsilon_{t-1}^{(i)}$. and this possibility is $(\frac{1}{2}\frac{1}{2^{k-i}})^2$, combining with the non-zero expectation value,
we have $\left(\sum_{i=1}^{k}(\frac{1}{2}\frac{1}{2^{k-i}})^2\right)\cdot\left[ln\left(m_0\right)-ln\left(2-m_0\right)\right]^4$.

(2) $\left(\varepsilon_{t+1}^{(j)}-\varepsilon_{t}^{(j)}\right)^{2}\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)^{2}$, which are non-zero for $i \neq j$, $\varepsilon_{t+1}^{(j)} \neq \varepsilon_{t}^{(j)}$ and $\varepsilon_{t}^{(i)} \neq \varepsilon_{t-1}^{(i)}$, the probability of its occurrence is $\sum_{i=1}^{k}\left(\frac{1}{2^{k-i}}\sum_{j=1,j\neq i}^{k}\frac{1}{2^{k-j}}\right)$.

Putting together these two possible forms we get
$\left[ln\left(m_0\right)-ln\left(2-m_0\right)\right]^4\cdot\left(\sum_{i=1}^{k}\frac{1}{2}\frac{1}{2^{k-i}}\sum_{j=1}^{k}\frac{1}{2}\frac{1}{2^{k-j}}\right)$.

(3) Form $\left(\varepsilon_{t+1}^{(j)}-\varepsilon_{t}^{(j)}\right)\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)\left(\varepsilon_{t}^{(j)}-\varepsilon_{t-1}^{(j)}\right)\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)$, which for $i \neq j$ and $\varepsilon_{t+1}^{(n)} \neq \varepsilon_{t}^{(n)} \neq \varepsilon_{t-1}^{(n)}$, $n = i, j$ are non-zero, and which implies
$2\left\{\sum_{i=1}^{k}\left((\frac{1}{2^{k-i}})^2\sum_{j=1,j\neq i}^{k}(\frac{1}{2^{k-j}})^2\right)\right\}\cdot\left[ln\left(m_0\right)-ln\left(2-m_0\right)\right]^4$.

Then we have the solution for the first component in the above moment condition:

$$E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})\right)^{2}\cdot\left(\sum_{i=1}^{k}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})\right)^{2}\right]$$
$$=\left[ln\left(m_0\right)-ln\left(2-m_0\right)\right]^4\left[\sum_{i=1}^{k}\frac{1}{2}\frac{1}{2^{k-i}}\sum_{j=1}^{k}\frac{1}{2}\frac{1}{2^{k-j}}+2\sum_{i=1}^{k}(\frac{1}{2}\frac{1}{2^{k-i}})^2\sum_{j=1,j\neq i}^{k}(\frac{1}{2}\frac{1}{2^{k-j}})^2\right]$$

The other components can be solved by recalling previous calculations. All in all, we finally arrive at:

$$E[X_{t+1,1}^2 \cdot Y_{t,1}^2]$$

$$= [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \tfrac{1}{16} \left[\sum_{i=1}^{k} \tfrac{1}{2} \tfrac{1}{2^{k-i}} \sum_{j=1}^{k} \tfrac{1}{2} \tfrac{1}{2^{k-j}} + 2 \sum_{i=1}^{k} (\tfrac{1}{2} \tfrac{1}{2^{k-i}})^2 \sum_{j=1, j\neq i}^{k} (\tfrac{1}{2} \tfrac{1}{2^{k-j}})^2 \right]$$

$$+ [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \tfrac{1}{16} \sum_{i=k+1}^{n} \tfrac{1}{2} \tfrac{1}{2^{n-i}} \sum_{i=k+1}^{n} \tfrac{1}{2} \tfrac{1}{2^{n-i}}$$

$$+ \tfrac{1}{8} [ln\,(m_0) - ln\,(2-m_0)]^4 \sum_{i=1}^{k} \tfrac{1}{2} \tfrac{1}{2^{k-i}} \sum_{i=k+1}^{n} \tfrac{1}{2} \tfrac{1}{2^{n-i}}$$

$$+ \left(E[ln|u_t|^2] - E[ln|u_t|]^2\right) \cdot [ln\,(m_0) - ln\,(2-m_0)]^2 \cdot \left(\sum_{i=1}^{k} \tfrac{1}{2} \tfrac{1}{2^{k-i}} + \sum_{i=k+1}^{n} \tfrac{1}{2} \tfrac{1}{2^{n-i}} \right)$$

$$+ E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 4E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|] + 4E[ln|u_{1,t}| \cdot ln|u_{2,t}|] E[ln|u_t|]^2$$

$$+ 3E[ln|u_t|^2]^2 - 4E[ln|u_t|^2] E[ln|u_t|]^2. \tag{A6}$$

$$E[X_{t+1,1}^2 \cdot X_{t,1}^2]$$

$$= E \left\{ \left[\tfrac{1}{2} \sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) + \tfrac{1}{2} \sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) + (ln|u_{1,t+1}| - ln|u_{1,t}|) \right]^2 \cdot \right.$$
$$\left. \left[\tfrac{1}{2} \sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \tfrac{1}{2} \sum_{l=k+1}^{n} (\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}) + (ln|u_{1,t}| - ln|u_{1,t-1}|) \right]^2 \right\}$$

$$= \tfrac{1}{16} E \left[\left(\sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \cdot \left(\sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \right] + \tfrac{1}{16} E \left[\left(\sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \cdot \left(\sum_{l=k+1}^{n} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \right]$$

$$+ \tfrac{1}{16} E \left[\left(\sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \left(\sum_{l=k+1}^{n} (\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}) \right)^2 \right] + \tfrac{1}{16} E \left[\left(\sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \left(\sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) \right)^2 \right]$$

$$+ \tfrac{1}{4} \left\{ 2E \left[\left(\sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \right] + 2E \left[\left(\sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) \right)^2 \right] \right\} \cdot \left(2E[ln|u_t|^2] - 2E[ln|u_t|]^2 \right)$$

$$+ \tfrac{1}{16} \cdot 4E \left[\sum_{i=1}^{k} \left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)} \right) \sum_{i=1}^{k} \left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)} \right) \right] E \left[\sum_{l=k+1}^{n} \left(\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)} \right) \sum_{l=k+1}^{n} \left(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)} \right) \right]$$

$$+ \tfrac{1}{4} \cdot 4E \left[\sum_{i=1}^{k} \left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)} \right) \sum_{i=1}^{k} \left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)} \right) \right] \cdot \left(E[ln|u_t|]^2 - E[ln|u_t|^2] \right)$$

$$+ \tfrac{1}{4} \cdot 4E \left[\sum_{l=k+1}^{n} \left(\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)} \right) \sum_{l=k+1}^{n} \left(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)} \right) \right] \cdot \left(E[ln|u_t|]^2 - E[ln|u_t|^2] \right)$$

$$+ 3E[ln|u_t|^2]^2 + E[ln|u_t|^4] - 4E[ln|u_t|^3] E[ln|u_t|].$$

(A7)

The first and second term are the same as the first one in the case $E[X_{t+1,1}^2, Y_{t,1}^2]$, and the rest are our familiars. Adding together, we have the result:

$$
\begin{aligned}
&E[X_{t+1,1}^2 \cdot X_{t,1}^2] \\
&= [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \tfrac{1}{16}\left[\sum_{i=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-i}} \sum_{j=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-j}} + 2\sum_{i=1}^{k} (\tfrac{1}{2}\tfrac{1}{2^{k-i}})^2 \sum_{j=1, j\neq i}^{k} (\tfrac{1}{2}\tfrac{1}{2^{k-j}})^2\right] \\
&+ [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \tfrac{1}{16}\left[\sum_{i=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{k-i}} \sum_{j=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{k-j}} + 2\sum_{i=k+1}^{n} (\tfrac{1}{2}\tfrac{1}{2^{k-i}})^2 \sum_{j=k+1, j\neq i}^{n} (\tfrac{1}{2}\tfrac{1}{2^{k-j}})^2\right] \\
&+ [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \tfrac{1}{8}\sum_{i=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-i}} \sum_{i=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{n-i}} \\
&+ \left(E[ln|u_t|^2] - E[ln|u_t|]^2\right) \cdot [ln\,(m_0) - ln\,(2-m_0)]^2 \cdot \left(\sum_{i=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-i}} + \sum_{i=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{n-i}}\right) \\
&+0.25\left[2ln\,(m_0)\,ln\,(2-m_0) - (ln\,(m_0))^2 - (ln\,(2-m_0))^2\right]^2 \left(\sum_{i=1}^{k} \left(\tfrac{1}{2}\tfrac{1}{2^{k-i}}\right)^2 \sum_{i=k+1}^{n} \left(\tfrac{1}{2}\tfrac{1}{2^{n-i}}\right)^2\right) \\
&+2\left[2ln\,(m_0)\,ln\,(2-m_0) - (ln\,(m_0))^2 - (ln\,(2-m_0))^2\right] \cdot \\
&\left(\sum_{i=1}^{k} \left(\tfrac{1}{2}\tfrac{1}{2^{k-i}}\right)^2 + \sum_{i=k+1}^{n} \left(\tfrac{1}{2}\tfrac{1}{2^{n-i}}\right)^2\right) \cdot \left(E[ln|u_t|]^2 - E[ln|u_t|^2]\right) \\
&+3E[ln|u_t|^2]^2 + E[ln|u_t|^4] - 4E[ln|u_t|^3]E[ln|u_t|].
\end{aligned}
$$

(A8)

B Lognormal case

$$cov[X_{t,1}, Y_{t,1}] = E[(X_{t,1} - E[X_{t,1}]) \cdot (Y_{t,1} - E[Y_{t,1}] = E[X_{t,1} \cdot Y_{t,1}]$$

$$= \tfrac{1}{4} E\left[\left(\sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right] + 2E[ln|u_{1,t}| \cdot ln|u_{2,t}|] - 2E[u_t]^2$$
$$= 0.5\sigma_\varepsilon^2 \sum_{i=1}^{k} \tfrac{1}{2^{k-i}} + 2E[ln|u_{1,t}| \cdot ln|u_{2,t}|] - 2E[u_t]^2. \tag{B1}$$

Because the non-zero outcomes occur when $\varepsilon_t^{(i)} \neq \varepsilon_{t-1}^{(i)}$, which implies:

$$(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})^2 = 2(E[(\varepsilon_t^{(i)})^2] - E[\varepsilon_t^{(i)}]^2) = 2\sigma_\varepsilon^2$$

$$cov[X_{t+1,1}, Y_{t,1}] = \tfrac{1}{4} E\left[\sum_{i=1}^{k} \left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}\right) \cdot \sum_{i=1}^{k} \left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right)\right] + E[u_t]^2 - E[ln|u_{1,t}| \cdot ln|u_{2,t}|]$$

$$= -0.25\sigma_\varepsilon^2 \sum_{i=1}^{k} \left(\tfrac{1}{2^{k-i}}\right)^2 + E[u_t]^2 - E[ln|u_{1,t}| \cdot ln|u_{2,t}|]. \tag{B2}$$

Because the non-zero outcomes occur when $\varepsilon_{t+1}^{(i)} \neq \varepsilon_t^{(i)} \neq \varepsilon_{t-1}^{(i)}$, which implies:

$$(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \cdot (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) = E[\varepsilon_t^{(i)}]^2 - E[(\varepsilon_t^{(i)})^2] = -\sigma_\varepsilon^2$$

$$cov[X_{t+1,1}, X_{t,1}]$$

$$= \tfrac{1}{4} E\left[\sum_{i=1}^{k} \left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}\right) \cdot \sum_{i=1}^{k} \left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}\right)\right] + \tfrac{1}{4} E\left[\sum_{l=k+1}^{n} \left(\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}\right) \cdot \sum_{l=k+1}^{n} \left(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}\right)\right]$$

$$+ E\left[ln|u_t|\right]^2 - E\left[ln|u_t|^2\right]$$

$$= -0.25\sigma_\varepsilon^2 \left[\sum_{i=1}^{k} \left(\tfrac{1}{2^{k-i}}\right)^2 \sum_{i=k+1}^{n} \left(\tfrac{1}{2^{n-i}}\right)^2\right] + E\left[ln|u_t|\right]^2 - E\left[ln|u_t|^2\right]. \tag{B3}$$

$$E[X_{t,1}^2 \cdot Y_{t,1}^2]$$

$$= \tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^4\right] + \tfrac{1}{16}E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\left(\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)})\right)^2\right]$$
$$+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\left(\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)})\right)^2\right]$$
$$+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\right]$$
$$+\tfrac{1}{4}\left\{2E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right] + 2E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\right]\right\}\cdot\left(2E[ln|u_t|^2] - 2E[ln|u_t|]^2\right)$$
$$+2E[(ln|u_{1,t}|)^2\cdot(ln|u_{2,t}|)^2] - 8E[(ln|u_{1,t}|)^2\cdot ln|u_{2,t}|]\cdot E[ln|u_t|]$$
$$+4E[ln|u_{1,t}|\cdot(ln|u_{2,t}|]^2 + 2E[(ln|u_t|)^2]^2$$

$$= 0.75\sigma_\varepsilon^4\sum_{i=1}^{k}\tfrac{1}{2^{k-i}} + 0.25\sigma_\varepsilon^4\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}\sum_{h=k+1}^{n}\tfrac{1}{2^{n-h}}$$
$$+0.25\sigma_\varepsilon^4\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{h=k+1}^{n}\tfrac{1}{2^{n-h}} + 0.25\sigma_\varepsilon^4\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}$$
$$+2\sigma_\varepsilon^2\left(E[ln|u_t|^2] - E[ln|u_t|]^2\right)\cdot\left(\sum_{i=1}^{k}\tfrac{1}{2^{k-i}} + \sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}\right)$$
$$+2E[(ln|u_{1,t}|)^2\cdot(ln|u_{2,t}|)^2] - 8E[(ln|u_{1,t}|)^2\cdot ln|u_{2,t}|]\cdot E[ln|u_t|]$$
$$+4E[ln|u_{1,t}|\cdot(ln|u_{2,t}|]^2 + 2E[(ln|u_t|)^2]^2. \tag{B4}$$

For the first term $E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^4\right]$, let's begin with $E\left[\left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right)^4\right]$, the non-zero value implies:

$$E\left[\left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right)^4\right] = 2E[\varepsilon_t^{(i)}]^4 + 6E[(\varepsilon_t^{(i)})^2]^2 - 8E[(\varepsilon_t^{(i)})^3]E[\varepsilon_t^{(i)}] = 12\sigma_\varepsilon^4.$$

This occurs with probability $2^{\frac{1}{k-i}}$. Then we have the solution:

$$E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^4\right] = 12\sigma_\varepsilon^4\cdot\sum_{i=1}^{k}\frac{1}{2^{k-i}}.$$

$E[X_{t+1,1}^2 \cdot Y_{t,1}^2]$

$$= \frac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)})\right)^2 \cdot \left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right]$$
$$+\frac{1}{16}E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)})\right)^2 \left(\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)})\right)^2\right]$$
$$+\frac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)})\right)^2 \left(\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)})\right)^2\right]$$
$$+\frac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2 \left(\sum_{l=k+1}^{n}(\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)})\right)^2\right]$$
$$+\frac{1}{4}\left\{2E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right] + 2E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\right]\right\} \cdot \left(2E[ln|u_t|^2] - 2E[ln|u_t|]^2\right)$$
$$+E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 4E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|] + 4E[ln|u_{1,t}| \cdot (ln|u_{2,t}|]E[ln|u_t|]^2$$
$$+3E[ln|u_t|^2]^2 - 4E[ln|u_t|^2]E[ln|u_t|]^2.$$

For the first term $E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)})\right)^2 \cdot \left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right]$, there are three different possible forms:

(1) $\left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}\right)^2 \left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right)^2$, has non-zero value only if $\varepsilon_{t+1}^{(i)} \neq \varepsilon_t^{(i)} \neq \varepsilon_{t-1}^{(i)}$. then $E\left[\left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}\right)^2 \left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right)^2\right] = E[\varepsilon_t^4] + 3E[\varepsilon_t^2]^2 - 4E[\varepsilon_t^3]E[\varepsilon_t] = 6\sigma_\varepsilon^4$. ($E[\varepsilon_t^3] = 3\lambda\sigma_\varepsilon^2 + \lambda^3$ and $E[\varepsilon_t^4] = 3\sigma_\varepsilon^4 + 6\lambda^2\sigma_\varepsilon^2 + \lambda^4$,) and the probability of this occurance is $(\frac{1}{2^{k-i}})^2$. Putting together we get $\left[\sum_{i=1}^{k}\left(\frac{1}{2^{k-i}}\right)^2\right] \cdot 6\sigma_\varepsilon^4$

(2) $\left(\varepsilon_{t+1}^{(j)} - \varepsilon_t^{(j)}\right)^2 \left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right)^2$, does not equal zero for $i \neq j$, $\varepsilon_{t+1}^{(j)} \neq \varepsilon_t^{(j)}$ and $\varepsilon_t^{(i)} \neq \varepsilon_{t-1}^{(i)}$. since $E\left[\left(\varepsilon_{t+1}^{(j)} - \varepsilon_t^{(j)}\right)^2 \left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right)^2\right] = 4E[(\varepsilon_t^{(i)})^2]^2 - 8E[(\varepsilon_t^{(i)})^2]E[\varepsilon_t^{(i)}]^2 + 4E[\varepsilon_t^{(i)}]^4 = 4\sigma_\varepsilon^4$, together with the probability, this overall contribution yields:
$\left[\sum_{i=1}^{k}\left(\frac{1}{2^{k-i}}\sum_{j=1,j\neq i}^{k}\frac{1}{2^{k-j}}\right)\right] \cdot 4\sigma_\varepsilon^4$

(3) $\left(\varepsilon_{t+1}^{(j)}-\varepsilon_{t}^{(j)}\right)\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)\left(\varepsilon_{t}^{(j)}-\varepsilon_{t-1}^{(j)}\right)\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)$, which for $i \neq j$ and $\varepsilon_{t+1}^{(n)} \neq \varepsilon_{t}^{(n)} \neq \varepsilon_{t-1}^{(n)}$, $n = i, j$ are non-zero, since $\left(\varepsilon_{t+1}^{(j)}-\varepsilon_{t}^{(j)}\right)\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)\left(\varepsilon_{t}^{(j)}-\varepsilon_{t-1}^{(j)}\right)\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right) = 4\sigma_{\varepsilon}^{4}$, we obtain a contribution $2\left[\sum_{i=1}^{k}\left(\frac{1}{2^{k-i}}\right)^2 \sum_{j=1,j\neq i}^{k}\left(\frac{1}{2^{k-j}}\right)^2\right] \cdot \sigma_{\varepsilon}^{4}$.

Combining those three cases, we have the result:

$$E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})\right)^2 \cdot \left(\sum_{i=1}^{k}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})\right)^2\right]$$

$$= 6\sigma_{\varepsilon}^{4} \cdot \sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^2 + 4\sigma_{\varepsilon}^{4} \cdot \sum_{i=1}^{k}\tfrac{1}{2^{k-i}} \sum_{j=1,j\neq i}^{k}\tfrac{1}{2^{k-j}} + 2\sigma_{\varepsilon}^{4} \cdot \sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^2 \sum_{j=1,j\neq i}^{k}(\tfrac{1}{2^{k-j}})^2$$

$$\begin{aligned}
&E[X_{t+1,1}^2 \cdot Y_{t,1}^2] \\
&= \tfrac{1}{16}\left[6\sigma_{\varepsilon}^{4} \cdot \sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^2 + 4\sigma_{\varepsilon}^{4} \cdot \sum_{i=1}^{k}\tfrac{1}{2^{k-i}} \sum_{j=1,j\neq i}^{k}\tfrac{1}{2^{k-j}} + 2\sigma_{\varepsilon}^{4} \cdot \sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^2 \sum_{j=1,j\neq i}^{k}(\tfrac{1}{2^{k-j}})^2\right] \\
&+0.25\sigma_{\varepsilon}^{4}\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}\sum_{h=k+1}^{n}\tfrac{1}{2^{n-h}} \\
&+0.25\sigma_{\varepsilon}^{4}\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{h=k+1}^{n}\tfrac{1}{2^{n-h}} + 0.25\sigma_{\varepsilon}^{4}\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}} \\
&+2\sigma_{\varepsilon}^{2} \cdot \left(E[ln|u_t|^2] - E[ln|u_t|]^2\right) \cdot \left(\sum_{i=1}^{k}\tfrac{1}{2^{k-i}} + \sum_{i=k+1}^{n}\tfrac{1}{2^{n-i}}\right) \\
&+E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 4E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|] + 4E[ln|u_{1,t}| \cdot (ln|u_{2,t}|]E[ln|u_t|]^2 \\
&+3E[ln|u_t|^2]^2 - 4E[ln|u_t|^2]E[ln|u_t|]^2
\end{aligned} \tag{B5}$$

$E[X_{t+1,1}^2 \cdot X_{t,1}^2]$

$$
\begin{aligned}
&= \tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})\right)^2\cdot\left(\sum_{i=1}^{k}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})\right)^2\right]+\tfrac{1}{16}E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})\right)^2\cdot\left(\sum_{l=k+1}^{n}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})\right)^2\right]\\
&+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})\right)^2\left(\sum_{l=k+1}^{n}(\varepsilon_{t}^{(l)}-\varepsilon_{t-1}^{(l)})\right)^2\right]+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})\right)^2\left(\sum_{l=k+1}^{n}(\varepsilon_{t+1}^{(l)}-\varepsilon_{t}^{(l)})\right)^2\right]\\
&+\tfrac{1}{4}\left\{2E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})\right)^2\right]+2E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_{t+1}^{(l)}-\varepsilon_{t}^{(l)})\right)^2\right]\right\}\cdot\big(2E[ln|u_t|^2]-2E[ln|u_t|]^2\big)\\
&+4\cdot\tfrac{1}{16}E\left[\sum_{i=1}^{k}\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)\sum_{i=1}^{k}\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)\right]E\left[\sum_{l=k+1}^{n}\left(\varepsilon_{t+1}^{(l)}-\varepsilon_{t}^{(l)}\right)\sum_{l=k+1}^{n}\left(\varepsilon_{t}^{(l)}-\varepsilon_{t-1}^{(l)}\right)\right]\\
&+4\cdot\tfrac{1}{4}E\left[\sum_{i=1}^{k}\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)\sum_{i=1}^{k}\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)\right]\cdot\big(E[ln|u_t|]^2-E[ln|u_t|^2]\big)\\
&+4\cdot\tfrac{1}{4}E\left[\sum_{l=k+1}^{n}\left(\varepsilon_{t+1}^{(l)}-\varepsilon_{t}^{(l)}\right)\sum_{l=k+1}^{n}\left(\varepsilon_{t}^{(l)}-\varepsilon_{t-1}^{(l)}\right)\right]\cdot\big(E[ln|u_t|]^2-E[ln|u_t|^2]\big)\\
&+3E[ln|u_t|^2]^2+E[ln|u_t|^4]-4E[ln|u_t|^3]E[ln|u_t|]\\
&= \tfrac{1}{16}\left[6\sigma_\varepsilon^4\cdot\sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^2+4\sigma_\varepsilon^4\cdot\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{j=1,j\neq i}^{k}\tfrac{1}{2^{k-j}}+2\sigma_\varepsilon^4\cdot\sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^2\sum_{j=1,j\neq i}^{k}(\tfrac{1}{2^{k-j}})^2\right]\\
&+\tfrac{1}{16}\left[6\sigma_\varepsilon^4\cdot\sum_{l=k+1}^{n}(\tfrac{1}{2^{n-l}})^2+4\sigma_\varepsilon^4\cdot\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}\sum_{j=k+1,j\neq i}^{n}\tfrac{1}{2^{n-j}}+2\sigma_\varepsilon^4\cdot\sum_{l=k+1}^{n}(\tfrac{1}{2^{n-l}})^2\sum_{j=k+1,j\neq i}^{n}(\tfrac{1}{2^{n-j}})^2\right]\\
&+0.25\sigma_\varepsilon^4\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}+0.25\sigma_\varepsilon^4\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\\
&+2\sigma_\varepsilon^2\left(\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}+\sum_{i=k+1}^{n}\tfrac{1}{2^{n-i}}\right)\cdot\big(E[ln|u_t|^2]-E[ln|u_t|]^2\big)\\
&+0.25\sigma_\varepsilon^4\sum_{i=1}^{k}\big(\tfrac{1}{2^{k-i}}\big)^2\cdot\sum_{i=k+1}^{n}\big(\tfrac{1}{2^{n-i}}\big)^2\\
&-\sigma_\varepsilon^2\cdot\left(\sum_{i=1}^{k}\big(\tfrac{1}{2^{k-i}}\big)^2+\sum_{i=k+1}^{n}\big(\tfrac{1}{2^{n-i}}\big)^2\right)\cdot\big(E[ln|u_t|]^2-E[ln|u_t|^2]\big)\\
&+3E[ln|u_t|^2]^2+E[ln|u_t|^4]-4E[ln|u_t|^3]E[ln|u_t|].
\end{aligned}
\tag{B6}
$$

Because the first term is identical with the first one of case $E[X_{t+1,1}^2, Y_{t,1}^2]$.

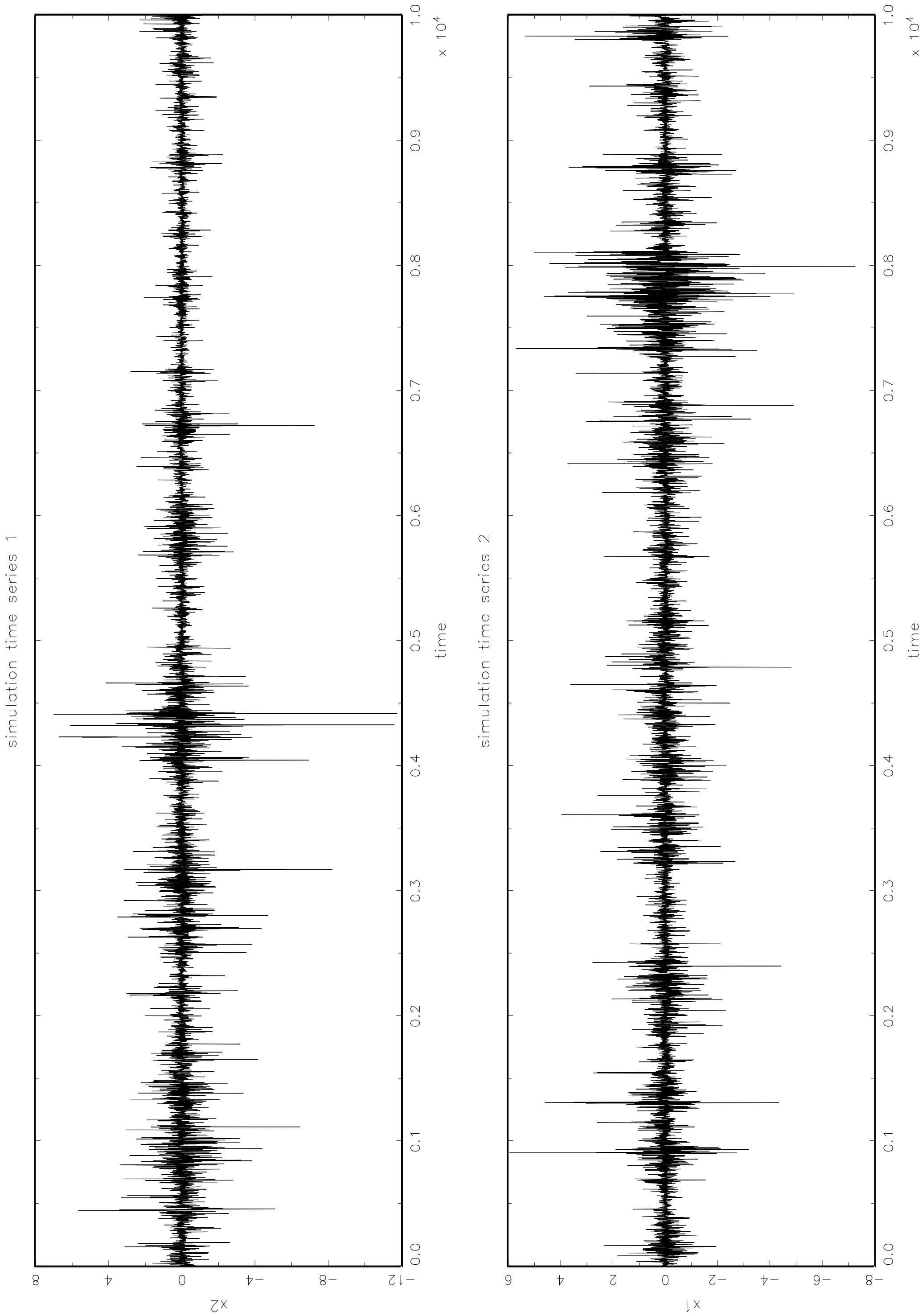

Figure 1: Simulation of the Bivariate Binomial Multi-Fractal Model with $m_0 = 1.4$, $\rho = 0.3$.

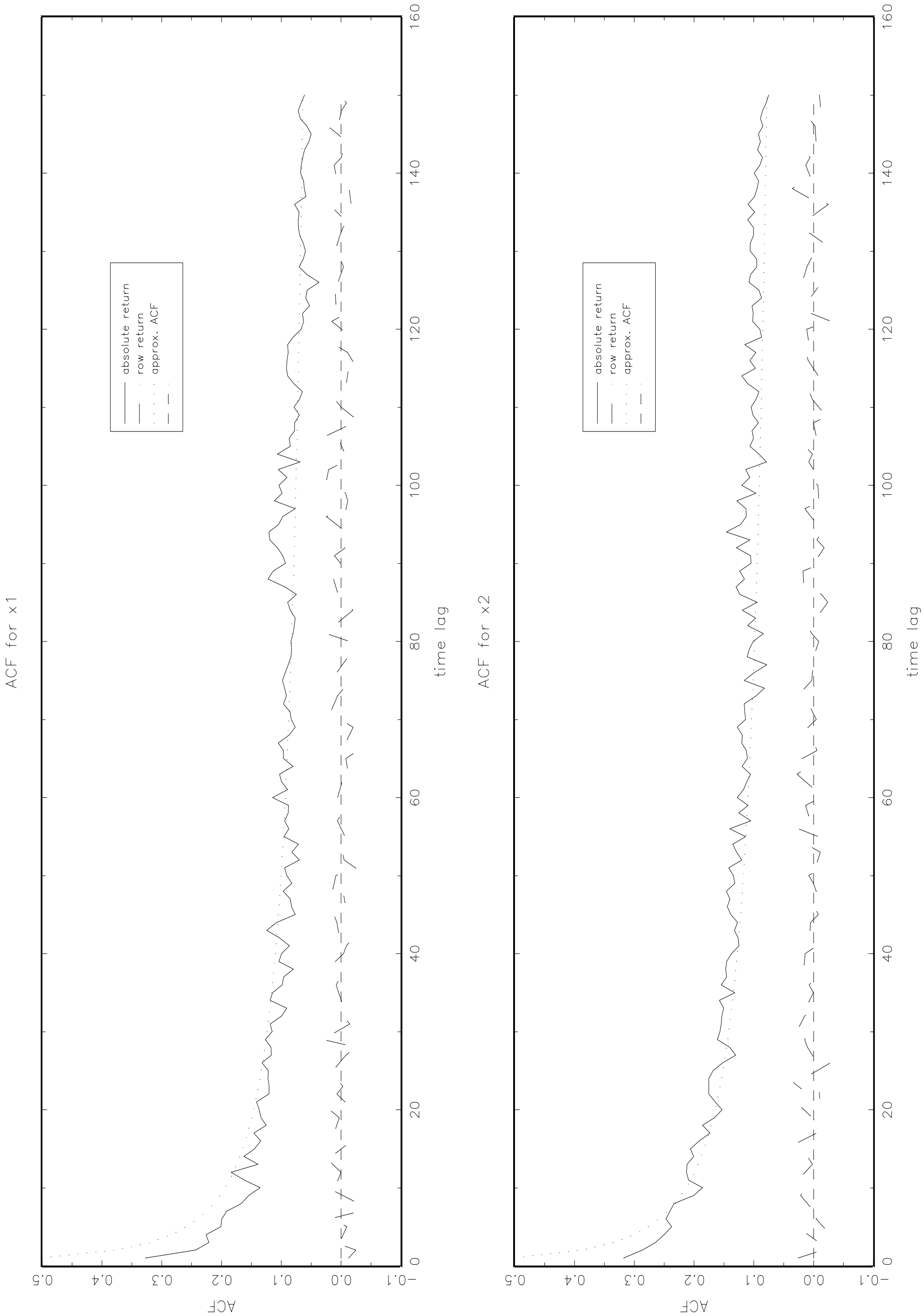

Figure 2: ACF for the Simulation of the Bivariate Binomial Multi-Fractal Model above.

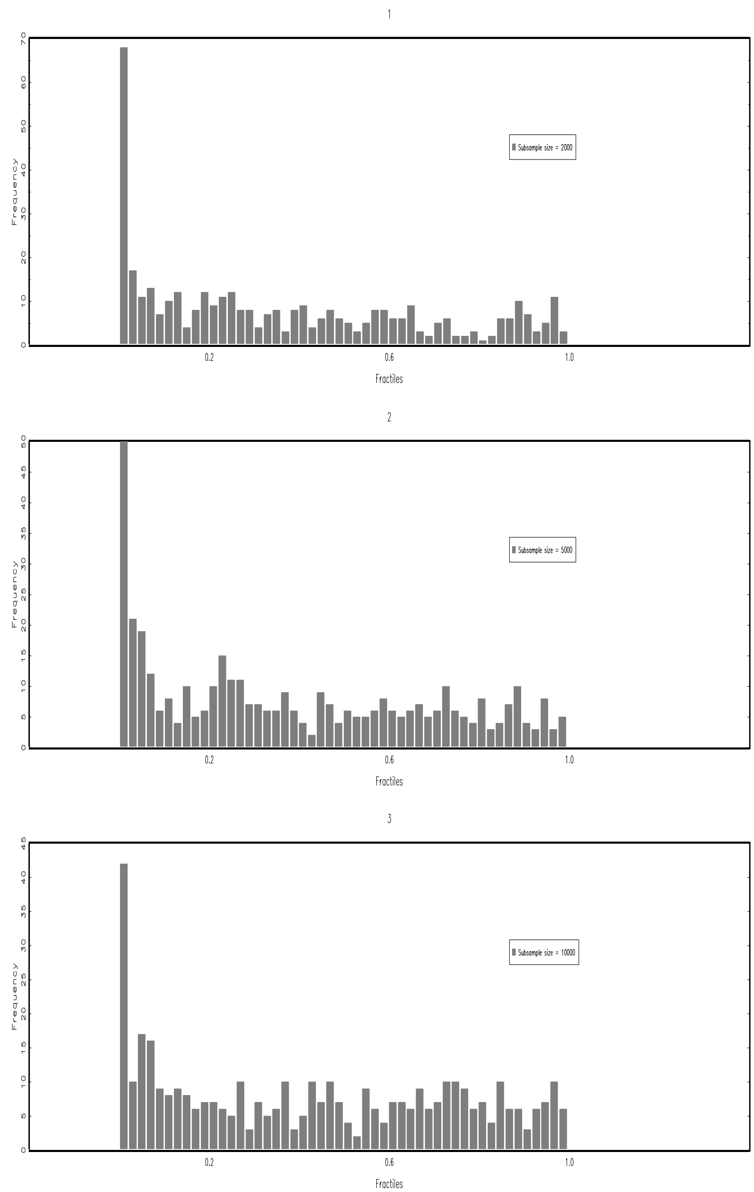

Figure 3: The distribution of p value for the test of overindentification restrictions for Binomial BMF. Three figures from up to down corresponding to three different sample size: $N_1 = 2,000$, $N_2 = 5,000$ and $N_3 = 10,000$.

Table 1: GMM Estimation of the Bivariate MF Binomial Model

		$k=4$						$k=8$					
		$\hat{m}_0$			$\hat{\rho}$			$\hat{m}_0$			$\hat{\rho}$		
		Bias	*SD*	*RMSE*	*Bias*	*SD*	*RMSE*	*Bias*	*SD*	*RMSE*	*Bias*	*SD*	*RMSE*
$m_0 = 1.20$	N_1	-0.0978	0.1209	0.1403	-0.0046	0.0807	0.0810	-0.0944	0.1330	0.1339	-0.0048	0.0834	0.0838
	N_2	-0.0590	0.1068	0.1208	-0.0061	0.0579	0.0652	-0.0660	0.1177	0.1248	-0.0018	0.0606	0.0605
	N_3	-0.0305	0.0993	0.1018	-0.0037	0.0372	0.0373	-0.0437	0.1002	0.1093	-0.0008	0.0319	0.0319
$m_0 = 1.30$	N_1	-0.1123	0.1257	0.1394	0.0056	0.0846	0.0847	-0.0891	0.1256	0.1364	0.0064	0.0833	0.0835
	N_2	-0.0042	0.1035	0.1117	0.0033	0.0589	0.0587	-0.0370	0.0889	0.0103	0.0019	0.0476	0.0476
	N_3	-0.0076	0.0568	0.0578	-0.0061	0.0389	0.0394	-0.0174	0.0689	0.0694	-0.0011	0.0358	0.0359
$m_0 = 1.40$	N_1	-0.0715	0.1201	0.1233	0.0050	0.0931	0.0935	-0.0606	0.1141	0.1297	0.0035	0.0910	0.0919
	N_2	-0.0158	0.0551	0.0562	-0.0081	0.0506	0.0520	-0.0122	0.0487	0.0520	-0.0072	0.0617	0.0621
	N_3	0.0027	0.0363	0.0375	-0.0097	0.0340	0.0339	-0.0027	0.0356	0.0359	-0.0079	0.0357	0.0365
$m_0 = 1.50$	N_1	-0.0409	0.0818	0.0827	-0.0091	0.0948	0.0951	-0.0315	0.0707	0.0773	-0.0092	0.0980	0.0983
	N_2	-0.0068	0.0374	0.0379	-0.0168	0.0569	0.0605	-0.0041	0.0336	0.0338	-0.0173	0.0595	0.0619
	N_3	0.0053	0.0263	0.0268	-0.0091	0.0379	0.0390	0.0019	0.0249	0.0250	-0.0093	0.0384	0.0429

Note: All simulations are based on the bivariate Multi-fractal process with the whole number of cascade levels equal to 20, $\rho = 0.5$, and eight moment conditions as in the Appendix A are used. Sample lengths are $N_1 = 2,000$, $N_2 = 5,000$ and $N_3 = 10,000$. Bias denotes the distance between the given and estimated parameter value, SD and RMSE denote the standard deviation and root mean squared error, respectively. For each scenario, 400 Monte Carlo simulations have been carried out.

Table 2: GMM Estimation of the Bivariate MF Lognormal Model

		$k=4$						$k=8$					
		$\hat{\lambda}$			$\hat{\rho}$			$\hat{\lambda}$			$\hat{\rho}$		
		Bias	*SD*	*RMSE*	*Bias*	*SD*	*RMSE*	*Bias*	*SD*	*RMSE*	*Bias*	*SD*	*RMSE*
$\lambda = 1.10$	N_1	-0.0234	0.0538	0.0586	0.0106	0.0916	0.0921	-0.0171	0.0573	0.0587	-0.0015	0.0937	0.0957
	N_2	-0.0069	0.0347	0.0353	0.0026	0.0494	0.0495	-0.0085	0.0359	0.0369	0.0038	0.0568	0.0569
	N_3	-0.0012	0.0243	0.0243	-0.0030	0.0383	0.0384	-0.0035	0.0242	0.0244	-0.0022	0.0377	0.0378
$\lambda = 1.20$	N_1	-0.0272	0.0606	0.0663	0.0063	0.0985	0.0986	-0.0251	0.0599	0.0649	0.0107	0.0966	0.0971
	N_2	-0.0071	0.0380	0.0387	-0.0042	0.0563	0.0564	-0.0091	0.0367	0.0377	0.0027	0.0655	0.0654
	N_3	-0.0022	0.0261	0.0263	-0.0008	0.0358	0.0360	-0.0011	0.0260	0.0262	-0.0035	0.0391	0.0395
$\lambda = 1.30$	N_1	-0.0262	0.0639	0.0650	0.0138	0.0950	0.0953	-0.0244	0.0610	0.0656	0.0035	0.0912	0.0913
	N_2	-0.0088	0.0397	0.0401	0.0054	0.0579	0.0576	-0.0070	0.0372	0.0378	-0.0022	0.0578	0.0587
	N_3	-0.0013	0.0265	0.0275	-0.0045	0.0362	0.0374	-0.0024	0.0268	0.0269	-0.0041	0.0389	0.0392
$\lambda = 1.40$	N_1	-0.0354	0.0650	0.0740	0.0085	0.0951	0.0953	-0.0313	0.0638	0.0710	0.0103	0.0897	0.0916
	N_2	-0.0054	0.0411	0.0414	-0.0048	0.0631	0.0631	-0.0080	0.0379	0.0385	-0.0031	0.0583	0.0582
	N_3	-0.0032	0.0270	0.0272	-0.0033	0.0398	0.0399	-0.0010	0.0269	0.0272	-0.0048	0.0396	0.0403

Note: All simulations are based on the bivariate Multi-fractal process with the whole number of cascade levels equal to 20, $\rho = 0.5$, and eight moment conditions as in the Appendix B are used. Sample lengths are $N_1 = 2,000$, $N_2 = 5,000$ and $N_3 = 10,000$. Bias denotes the distance between the given and estimated parameter value, SD and RMSE denote standard deviation and root mean squared error, respectively. For each scenario, 400 Monte Carlo simulations have been carried out.

Table 3: Comparison of ML and GMM Estimations

		ML						GMM					
		$\hat{m}_0$			$\hat{\rho}$			$\hat{m}_0$			$\hat{\rho}$		
		$Bias$	SD	$RMSE$	$Bias$	SD	$RMSE$	$Bias$	SD	$RMSE$	$Bias$	SD	$RMSE$
$m_0 = 1.20$	N_1	-0.0088	0.0191	0.0209	0.0070	0.0157	0.0171	-0.0793	0.1424	0.1628	-0.0104	0.0821	0.0826
	N_2	-0.0101	0.0116	0.0153	0.0070	0.0103	0.0125	-0.0801	0.1236	0.1472	-0.0022	0.0462	0.0465
	N_3	-0.0087	0.0088	0.0123	0.0082	0.0069	0.0107	-0.0487	0.1068	0.1177	-0.0010	0.0334	0.0334
$m_0 = 1.30$	N_1	-0.0093	0.0175	0.0197	0.0162	0.0180	0.0242	-0.0993	0.1542	0.1687	-0.0014	0.0831	0.0832
	N_2	-0.0109	0.0096	0.0145	0.0129	0.0114	0.0172	-0.0595	0.1188	0.1327	0.0014	0.0450	0.0450
	N_3	-0.0130	0.0060	0.0143	0.0135	0.0077	0.0156	-0.0259	0.0839	0.0877	-0.0007	0.0342	0.0342
$m_0 = 1.40$	N_1	-0.0113	0.0156	0.0192	0.0217	0.0202	0.0296	-0.0679	0.1324	0.1487	0.0015	0.0833	0.0832
	N_2	-0.0133	0.0091	0.0161	0.0236	0.0109	0.0259	-0.0313	0.0871	0.0925	-0.0013	0.0479	0.0479
	N_3	-0.0123	0.0064	0.0139	0.0234	0.0093	0.0253	-0.0061	0.0402	0.0406	-0.0050	0.0347	0.0351
$m_0 = 1.50$	N_1	-0.0127	0.0103	0.0163	0.0286	0.0202	0.0349	-0.0353	0.0796	0.0870	-0.0062	0.0901	0.0902
	N_2	-0.0110	0.0092	0.0143	0.0287	0.0130	0.0315	-0.0117	0.0416	0.0432	-0.0083	0.0493	0.0500
	N_3	-0.0114	0.0054	0.0126	0.0306	0.0084	0.0317	-0.0008	0.0278	0.0278	-0.0110	0.0369	0.0384

Note: Simulations are based on the Bivariate Binomial Multi-Fractal process with $n = 5$, $k = 2$, which is almost the limit of computational feasibility. We fixed initial $\rho = 0.5$, and sample sizes are $N_1 = 2,000$, $N_2 = 5,000$ and $N_3 = 10,000$. Bias denotes the distance between the given and estimated parameter value, SD and RMSE denote the standard deviation and root mean squared error, respectively. For each scenario, 400 Monte Carlo simulations have been carried out.

Table 4: ML estimation

$\hat{\theta}$	Sub-sample Size	*Bias*	*SD*	*RMSE*
	N_1	-0.0106	0.0170	0.0200
$\hat{m_0}$	N_2	-0.0096	0.0118	0.0152
	N_3	-0.0101	0.0075	0.0126
	N_1	-0.0016	0.0222	0.0221
$\hat{\sigma_1}$	N_2	-0.0004	0.0140	0.0140
	N_3	0.0002	0.0086	0.0086
	N_1	-0.0033	0.0227	0.0228
$\hat{\sigma_2}$	N_2	-0.0017	0.0128	0.0128
	N_3	-0.0010	0.0087	0.0087
	N_1	0.0099	0.0208	0.0230
$\hat{\rho}$	N_2	0.0108	0.0123	0.0163
	N_3	0.0110	0.0079	0.0135

Note: Simulations are based on the Bivariate Binomial Multi-Fractal process with $n = 5$, $k = 2$, which is almost the limit of computational feasibility, and initial value $m_0 = 1.3$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\rho = 0.5$. Sample lengths are $N_1 = 2,000$, $N_2 = 5,000$ and $N_3 = 10,000$. Bias denotes the distance between the given and estimated parameter value, SD and RMSE denote the standard deviation and root mean squared error, respectively. For each scenario, 400 Monte Carlo simulations have been carried out.

Table 5: Failure rates for multi-period Value-at-Risk forecasts

		One day horizon			Two days horizon			Five days horizon		
		DOW	*NIK*	*EW*	*DOW*	*NIK*	*EW*	*DOW*	*NIK*	*EW*
Stocks	$p = 10\%$	0.1074	0.0896	0.0974	0.1025	0.0912	0.1062	0.1052	0.0872	0.1110
	$p = 5\%$	0.0625^{+}	0.0422	0.0488	0.0627^{+}	0.0451	0.0584	0.0600	0.0453	0.0645^{+}
	$p = 1\%$	0.0180	0.0081	0.0130	0.0193	0.0100	0.0165	0.0186	0.0087	0.0191
		BP	*AUD*	*EW*	*BP*	*AUD*	*EW*	*BP*	*AUD*	*EW*
FXs	$p = 10\%$	0.1059	0.0966	0.1104	0.1088	0.1035	0.1108	0.1051	0.0926	0.1051
	$p = 5\%$	0.0542	0.0487	0.0512	0.0547	0.0501	0.0557	0.0563	0.0419	0.0409
	$p = 1\%$	0.0106	0.0077	0.0085	0.0097	0.0092	0.0100	0.0118	0.0068	0.0069
		T1	*T2*	*EW*	*T1*	*T2*	*EW*	*T1*	*T2*	*EW*
Bonds	$p = 10\%$	0.1010	0.1104	0.1083	0.1043	0.1096	0.1070	0.0902	0.1063	0.1041
	$p = 5\%$	0.0577	0.0665^{+}	0.0642^{+}	0.0530	0.0659^{+}	0.0662^{+}	0.0541	0.0633^{+}	0.0535
	$p = 1\%$	0.0156	0.0187	0.0191	0.0149	0.0263^{+}	0.0237^{+}	0.0175	0.0191	0.0227^{+}

Note: This table shows the failure rate (ratio of the observations above the VaR). Stocks are Dow Jones Composite 65 Average Index and NIKKEI 225 Stock Average Index (NIK), FXs are Foreign Exchange rate of British Pound (BP) and Australian Dollar (AUD) to US Dollar, Bonds are the US 1-Year and 2-Year Treasury Constant Maturity Rate (T1, T2 respectively). Half of data for GMM estimation, which is based on the BMF Binomial model and BMF Lognormal model using eight moment conditions as in the Appendix, half out-sample data for VaR forecasting and failure rate calculation. EW denotes Equal-Weight portfolio. $+$ and $*$ denote too risky and too conservative VaR, with threshold $10\% \pm 0.02$, $5\% \pm 0.01$, and $[0.005, 0.02]$ for 1% level.

Table 6: Failure rates for multi-period Value-at-Risk forecasts

		One day horizon			Two days horizon			Five days horizon		
		DOW	*NIK*	*EW*	*DOW*	*NIK*	*EW*	*DOW*	*NIK*	*EW*
Stocks	$p = 10\%$	0.1175	0.0959	0.0904	0.1140	0.0968	0.0967	0.1125	0.0898	0.1000
	$p = 5\%$	0.0690^{+}	0.0439	0.0435	0.0677^{+}	0.0465	0.0495	0.0650^{+}	0.0475	0.0563
	$p = 1\%$	0.0173	0.0063	0.0082	0.0180	0.0098	0.0120	0.0191	0.0088	0.0162
		BP	*AUD*	*EW*	*BP*	*AUD*	*EW*	*BP*	*AUD*	*EW*
FXs	$p = 10\%$	0.1176	0.1119	0.0903	01168	0.1143	0.0875	0.1095	0.0951	0.0892
	$p = 5\%$	0.0577	0.0516	0.0403	0.0560	0.0517	0.0302^{*}	0.0563	0.0431	0.0244^{*}
	$p = 1\%$	0.0084	0.0054	0.0031^{*}	0.0082	0.0070	0.0027^{*}	0.0106	0.0056	0.0031^{*}
		T1	*T2*	*EW*	*T1*	*T2*	*EW*	*T1*	*T2*	*EW*
Bonds	$p = 10\%$	0.1001	0.1092	0.1060	0.1020	0.1085	0.1041	0.0872	0.1019	0.0982
	$p = 5\%$	0.0572	0.0662^{+}	0.0611	0.0521	0.0648^{+}	0.0653^{+}	0.0527	0.0615	0.0528
	$p = 1\%$	0.0158	0.0189	0.0191	0.0154	0.0273^{+}	0.025^{+}	0.0178	0.0230^{+}	0.0241^{+}

Note: This table shows the failure rate (ratio of the observations above the VaR). Stocks are Dow Jones Composite 65 Average Index and NIKKEI 225 Stock Average Index (NIK), FXs are Foreign Exchange rate of British Pound (BP) and Australian Dollar (AUD) to US Dollar, Bonds are the US 1-Year and 2-Year Treasury Constant Maturity Rate (T1, T2 respectively). Half of data for ML estimation, which is based on the BMF binomial model, half out-sample data for VaR forecasting and failure rate calculation. EW denotes Equal-Weight portfolio. $+$ and $*$ denote too risky and too conservative VaR, with threshold 10% $\pm$ 0.02, 5% $\pm$ 0.01, and [0.005, 0.02] for 1% level.